DOCUMENT RESUME

ED 312 229 SP 031 526

AUTHOR Ayers, Jerry B.; And Others
TITLE Follow-Up Evaluation of Graduates of the Teacher Education Programs of Tennessee Technological University.
INSTITUTION Tennessee Technological Univ., Cookeville. Center for Techer Education Evaluation.
PUB DATE 30 Jun 89
NOTE 54p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Authoritarianism; *Beginning Teachers; *Followup Studies; *Graduate Surveys; Higher Education; *Program Evaluation; Teacher Education Programs; *Test Validity
IDENTIFIERS American College Testing Program; Classroom Observations Keyed Effectiveness Res; Kentucky Observation Instrument; Pre Professional Skills Tests; *Tennessee Technological University

ABSTRACT

During 1987-87 through 1988-89, longitudinal follow-up studies were conducted of the graduates of the teacher education programs at Tennessee Technological University. The sample consisted of 60 first-year teachers who had completed their bachelor's degrees. During 1986-87 the Classroom Observation Keyed for Effectiveness Research was utilized as the primary research instrument. In 1987-88 and again in 1988-89, the Kentucky Observation Instrument was utilized. This document combines the results of these 3 years of activities into one report. The first four chapters center on the follow-up activities that were conducted during the 3-year period. Chapter five examines the relationships between the scores from the American College Test, the Pre-Professional Skills Test, the College Outcomes Measures Project test of the American College Testing Program and the National Teacher Examinations and the overall grade-point-average of the students required to take these tests. The final chapter presents the results of a study of the level of anti-democratic tendencies (authoritarianism) exhibited by groups of beginning teachers who graduated from the Tennessee Tech University teacher education programs. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED312229

Report 89-3-6

Follow-up Evaluation of Graduates of the Teacher Education Programs of Tennessee Technological University

Report Compiled By

Jerry B. Ayers, Director

Center for Teacher Education Evaluation
Tennessee Technological University
Cookeville, Tennessee 38505

June 30, 1989

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J.B. Ayers

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

[1]The evaluation was conducted as a part of the activities of the Center for Teacher Education Evaluation, Tennessee Technological University.

SP031526

BEST COPY AVAILABLE

TABLE OF CONTENTS

Preface..... iii

Chapter 1-Assessment of Teaching Performance and Classroom Observation 1987..... 1

Chapter 2-Assessment of Teaching Performance and Classroom Observation 1988..... 20

Chapter 3-Assessment of Teaching Performance and Classroom Observation 1989..... 29

Chapter 4-The NTE, PPST and Classroom Performance..... 33

Chapter 5-Relationship Between Scores from the ACT, PPST, NTE, ACTCOMP, and GPA..... 40

Chapter 6-Authoritarianism in Beginning Teachers..... 45

Preface

In 1973, the Office of the Assistant to the Dean and later the Office of the Associate Dean of the College of Education at Tennessee Technological University initiated a longitudinal study of the graduates of the teacher education programs of the University. The basic effort continued until the Summer of 1986. The project continued until 1986 utilizing essentially the same instrumentation. A number of reports and documents were prepared related to the longitudinal study. Many of the reports and documents have been disseminated in the literature, presented at professional meetings, and deposited in the collections of the Educational Resources Information Centers.

In 1986, based on the success of the longitudinal studies conducted at Tennessee Technological University, the State of Tennessee through the Tennessee Higher Education Commission, funded a Center of Excellence at the University. The Center for Teacher Education Evaluation was charged with the responsiblity of developing models and systems for the evaluation of teacher education programs. A secondary purpose of the Center was to conducted research related to program evaluation and to disseminate information about the results of the work that was conducted.

It seemed logical that the work of the Office of the Associate Dean should be combined with the activities of the Center. During 1986-7 through 1988-89, the longitudinal follow-up studies continued. However, experimental instruments and systems were utilized in an effort to make the project more meaningful and to provide, what was hoped, information that would be useful for program development at the University. During 1986-87 the Classroom Observation Keyed to Effectiveness Research was utilized as the primary research instrument. In 1987-88 and again in 1988-89 the Kentucky Observation Instrument was utilized.

This document combines the results of three years of activities into one report. This report is composed of four chapters that were prepared by the author and two that were prepared by graduate research assistants. The first four chapters center on the follow-up activities that were conducted during the three year period. Chapters V and VI reported the results of specialized studies that may be of interest to the reader.

Readers are invited to make comments and to request additional information about the process. Inquires can be addressed to the Office of the Associate Dean, College of Education, Tennessee Technological University, Cookeville, TN 38505.

Jerry B. Ayers,
Associate Dean and
Director, CTEE

Chapter I

ASSESSMENT OF TEACHING PERFORMANCE AND CLASSROOM OBSERVATION 1987

Jerry B. Ayers

Institutions of higher education must be accountable for the quality of the graduates of their teacher education programs. The key to innovation in the content and curriculum of teacher education programs is systematic evaluation. Tennessee Technological University has long recognized that improvement of curricula and programs is dependent upon an adequate system of evaluation.

In 1973, Tennessee Technological University developed and implemented a longitudinal follow-up study of the graduates of the teacher education programs. The purpose of the study was to gathered information for program change and improvement. The study continued utilizating the same basic instruments through 1985-86 (Ayers, 1986). During the 13 year period useful information was collected and the teacher preparation programs of the institution were improved. Since the inception of the longitudinal studies, the same basic instruments and sources of data were used to gather follow-up information. Each year data were collected from a sample of graduates in four basis areas using four sets of instruments as follows:

1. Personal information derived from transcripts and official University documents. Also each participant completed the California F-Scale Forms 45 and 40 [see Chapter VI].

2. Evaluations by principals and supervisors.

3. Evaluations by the classroom students of the graduates.

4. Through direct observation, by independent observers, in the classrooms of the graduates.

In 1986, the Center for Teacher Education Evaluation (CTEE) was established at the University. The specific major goals of the CTEE included:

1. Contributing to the continual improvement of teacher education programs through systematic development and testing of teacher education evaluation models.

2. Furthering research in the area of teacher education program evaluation.

3. Establishing improved delivery systems for information about teacher education program evaluation.

In the fall of 1986, the longitudinal follow-up studies of the teacher education programs were combined with the activities of the CTEE. It was felt that this would provide for a more efficient operation.

As a part of the original activities to develop the proposal for the establishment of the CTEE, an extensive review was made of the literature to locate models and relevant research for conducting follow-up evaluation studies. The Classroom Observation Keyed to Effectiveness Research (COKER) developed by Coker and Coker (1982) appeared to be an instrument that would be worthy of consideration as a primary research vehicle for gathering data for program improvement. The instrument was developed over an extended period of time and had many advantages as a device to gather useful information. Therefore, the decision was made to modify the original Tennessee Technological University Teacher Evaluation Model and develop a new system built around the COKER.

The remainder of this report was divided into four sections as follows: description of the COKER, methodology employed in the 14th year of the follow-up study, results of the study, and conclusions and recommendations for the future.

Description of the COKER

The Classroom Observation Keyed for Effectiveness Research (COKER) is a low-inference, direct observation instrument used to record the behaviors of pupils and their teacher in actual classrooms. This measurement-based evaluation of teacher performance is a relatively new approach to the evaluation of teachers (Medely, Coker & Soar, 1984).

There are many studies of teacher effectiveness that dictate the desirability of using teacher performance as a basis for individual and program evaluation decisions. This approach offered the promise as a basis for studying the evaluation of teacher education programs and providing empirical data to be used as the basis for change. The instrument was used on a limited basis in program evaluation work at such institutions as Georgia State University, the University of Toledo (Dickson & Wiersma, 1980), and at the University of Florida. The instrument or a modification was also used to evaluate teachers in such systems as the Humble ISD, TX; Volusia County, FL; and the State of Alabama.

The importance and usefulness of classroom observation systems are based on the assumption that what teachers do in the classroom affects pupil learning (Coker & Coker, 1982). Research is far from conclusive as to exactly what teacher behaviors affect learning, although there is evidence that some of the widely held beliefs may in fact be myths (Coker, Medley & Soar, 1980). In a synthesis of the research literature, (Medley (1977), concluded that the effects of many procedures and behaviors are situation specific or they depend upon factors associated with the pupils.

If one accepts this basic assumption, the primary question becomes: "What teacher behaviors contribute to teacher effectiveness?" Teacher effectiveness is not measured directly; but, the behaviors which characterize effective teachers are measured (Medley, et al., 1981; Coker & Coker, 1982).

The minimum conditions for valid evaluation of teacher performance includes for distinct steps, each of which must be performed with a high

degree of objectivity. The four steps, as applied to the evaluation of teacher performance are as follows:

1. Setting, defining, or agreeing upon a task to be performed;

2. Making a documentary, quantifiable record of the behavior of the teacher while the task is being performed;

3. Quantifying the record, that is, deriving a score or set of scores from it; and

4. Comparing the score with a predetermined standard.

The COKER is an instrument that was developed and refined over a period of almost a decade. Initial work began in 1973 with the current version of the instrument being established in 1982. The COKER was derived from careful analyses of five instruments that were developed and validated. The five observations that served as a basis for the development of the COKER included: OScAR 5V: Observation Schedule and Record, Form 5 Verbal (Medley, 1973); STARS: Spaulding Teacher Activity Recording Schedule (Spaulding, 1976); FLACCS: Florida Climate and Control System (Soar, Soar & Rogosta, 1971); TPOR: Teacher Practices Observation Record (Brown, 1970); and CASES: Coping Analysis Schedule for Educational Settings (Spaulding, 1976). The rationale for synthesizing items from the five instruments was to capture the breadth of the original devices. Selection of items was based on correlations between scores on individual items and pupil gain and the collective judgements of experts (Coker & Coker, 1982).

No attempt will be made at this point to fully describe all of the various parts of the COKER. The reader is referred to the Manual that accompanies the instrument (Coker & Coker, 1982). Basically the COKER consists of 389 cells in two sections. One section focuses on the observed interaction in the classroom, while the second section deals with student teacher affect. Trained observers sample teacher behaviors in five minute segments over an extended period of several months. Usually up to six observations are made. A score sheet is completed during each observation (see Appendix to this paper).

Through a process of factor analysis (Medely, et al, 1981; Coker & Coker, 1982; Wiersma, 1986), scores can be derived for performance objectives that have been established for a given program. The Appendix contains a set of objectives that were developed for the Humble Independent School District in Texas (Say, Georgiades, Porter & Goldstein, 1987). Scores can be derived for each teacher for each of the 13 major objectives. Two systems of scoring can be used with the COKER. Relative z-scores (mean of 50 and a standard deviation of 10) can be derived for a given group of teachers for each objective. This in turn allows for the development a comprehensive profile of each teacher and also allows for comparison of teachers across classes. The Appendix contains profiles for two graduate of the teacher preparation programs of Tennessee Technological University keyed to the objectives of the Humble ISD. One teacher was functioning above the mean on all of the objectives, while the scores achieved on the objectives by the second teacher

varied from over one standard deviation above the mean to one standard deviation below the mean. This system allows for the placement of teachers and provides direct feedback to the teachers on areas of weakness in comparison to other teachers. This scoring system has little value for determining weaknesses in the teacher preparation programs of the University.

A second scoring system for the COKER, called the "proportionality score" system provides for the computation of a score from 0 to 1.0 for each objective for each teacher. In turn absolute scores can be computed for each objective and comparisons are possible with other groups of teachers and weaknesses can be noted in the programs of the group.

The COKER is a difficult instrument to score. The major Center for scoring the instrument is located at Georgia State University in Atlanta. In the past year two other locations have begun some limited scoring of the instrument, however, the costs are excessive.

Methodology Employed at Tennessee Technological University

All members of the staff of the Center for Teacher Education Evaluation participated in a three day training program conducted by Dr. Homer Coker (the primary author of the instrument). This 30 hour program helped initially prepare eight individuals who were directly responsible for the collection of all data. In addition to the initial sessions with Dr. Coker the individuals trained to conduct the actual evaluation spent about 25 additional hours observing in classrooms and working with training tapes. The purpose of the extensive trainin[illegible]as to establish a high degree of reliability between the observers.

All individuals who had completed either a bachelor's or master's degree during 1985-86 and were known to be teaching in grades K-12 within a 75 miles radius of Cookeville, TN were asked to participate in the study. A radius of 75 miles was chosen because of the availability of funds to support travel. A total of 44 individuals who completed the Bachelor of Science degree and 28 who had completed a Master of Arts. Between approximately February 15, 1987 and May 15, 1987 each participant was visited on five or six occasion by a trained observer. Data were collected utilizing the COKER as the primary observation instrument. In addition principals were asked to complete a a supervisory rating device previously used in the follow-up studies (Ayers, 1986), personal data were collected from University records, and all participants completed the California F-Scale Forms 45 and 40.

The COKER was submitted for analysis to Georgia State University. After nearly a year it was discovered that the reliability of the observers was not at a high enough level to allow for scoring utilizing an absolute set of measures derived from the objectives of the teacher education program of Tennessee Technological University. A second attempt was made to score the instrument against the objectives of the Humble IDS. These attempts were unsuccessful. Scoring on a relative basis was possible (see Appendix). Therefore, mean scores for all objectives were "50."

This is a major limitation on the use of the COKER. As a result of these problems, the majority of the data for the 1987 evaluation program were lost. However, it was possible to report some results from the use of the other instrumentation and some results can be reported comparing the results from the various instruments or sources of data.

Results

The results section is divided into three parts with further subdivisions based on degree level (i.e., B.S. or M.A.). The first part reports personal data about the graduates, the second summarizes the results from the principal's evaluations, and the third examines the results of the use of the COKER.

Personal Information

The American College Test (ACT) is usually completed by individuals prior to initial admission to collegiate study. Table 1 shows a summary of the descriptive statistics available for the ACT scores for the individuals who participated in the study. The mean scores for the graduates are about comparable with those of other individuals who have participated in the study in past years. As a whole the mean scores of the gradutes are somewhat above those of all graduates from the teacher education programs.

Table 1

ACT Scores for Participants

ACT Test	B.S. (N=37)		M.A. (N=17)	
	Mean	SD	Mean	SD
English	19.0	5.1	19.5	3.4
Mathematics	15.2	7.0	19.4	6.2
Social Studies	18.2	6.7	18.7	5.4
Natural Science	21.3	5.9	20.2	4.6
Composite	18.7	5.8	19.6	3.5

Scores from the Core Battery of the National Teacher Examinations (NTE) were available for 34 bachelor's graduates. Table 2 shows the mean scores, standard deviation and approximate percentile for each of the three NTE tests. The subjects in the study had achieved scores at or above the 50th percentile on the NTE. Comparison of scores derived from the participants in this group with earlier groups was not possible since the NTE Core Battery was changed in 1985 and scores are not comparable.

Table 2

NTE Scores (N=34)

Test	$\bar{X}$	SD	Percentile
Communication Skills (CS)	663.3	11.1	53
General Knowledge (GK)	659.0	11.2	48
Professional Knowledge (PK)	660.9	12.2	53

The bachelor's level graduates had completed their programs of study with a mean grade point average of 3.18 (SD=0.42). The mean grade point average for master's level graduates of 3.70 (SD=0.31). These grade point averages are comparable with other groups of graduates who have participated in the longitudinal studies.

A total of 28 M.A. graduates had completed the Miller Analogies Test (MAT). The mean score was 37.1 (SD=14.8). This placed the graduates at approximately the 50th percentile when compared with national norms for the test.

The 44 graduates who had attained the bachelors' degree during 1985-86 and volunteered to participate in the study had completed a mean of slightly less than one year (0.9 years, SD=0.5) as teachers at the time of their initial entry into the study. In comparison the 28 individuals who had received master's degree had complete a mean of 6.8 years (SD=4.0) in the classroom.

The graduates in both groups completed the California F-scale Forms 45 and 40 developed by Adorno et al. (1950 to measure anti-democratic tendencies. The F-scale, in this study, was designed to serve as a measure of personality [see Chapter VI]. Mean F-scale scores for the two groups are shown in Table 3. The M.A. graduates achieved a higher score (not significant) than the B.S. graduates and thus appeared to be more authoritarian. This has been found in other studies of the graduates of the teacher education programs.

Table 3

F-scale Scores

Degree Level	N	$\bar{X}$	SD	t-ratio
Bachelor's	33	104.8	21.1	1.29
Master's	18	112.6	21.1	

In summary the graduates who participated in the study appeared to be comparable with other groups from previous years. Although inferential statistics were not employed, it appeared that the groups were achieving at slightly higher levels on standardized tests and they were generally more authoritarian in their beliefs.

Principal's Evaluations

The principal of each subject was asked to complete a four item instrument that was originally developed at Kansas State Teachers College (Sandefur, 1970). The instrument has been used in previous years of the study. Table 4 shows the results of the use of the instrument. There were no differences between the two groups. Mean scores were comparable with those from earlier years.

Table 4

Descriptive Statistics for Supervisory Evaluation Instrument

Factor	B.S. (N=40) $\bar{X}$	B.S. (N=40) SD	M.A. (N=24) $\bar{X}$	M.A. (N=24) SD
Subject Matter Competence	4.4	0.7	4.4	0.6
Relations With Students	4.5	0.6	4.5	0.7
Appropriateness of Assign.	4.4	0.7	4.4	0.6
Overall Effectiveness	4.4	0.7	4.5	0.6

Observation Instrument

Results obtained with the use of the COKER were disappointing. The information that was gathered could not be processed for use in program improvement. The profiles (see Appendix) were excellent for use by the individuals who had participated in the study.

It was possible to make some limited uses of the data gathered through observation. Correlations of z-scores were made with principal ratings and from the core battery tests of the NTE. Table 5 contains a summary of the characteristics from the Humble ISD that were used in this part of the study.

Data were available from a sample of 34 individuals who had completed the bachelor's degree and were in their first year of teaching. Table 6 shows a summary of the correlations between the 13 factors from the COKER and the NTE Core Battery scores and the overall principals rating.

Table 5

Summary of Characteristics Measured by the COKER

Focus-	Establishes focus of the lesson.
Org. Pres.-	Organizes classroom presentations.
Directions-	Gives clear and easily understood directions.
Participation-	Provides opportunities for students to actively and successfully participates.
Evaluates-	Evaluates student progress and provides feedback.
Org. Class.-	Provides organization in the classroom.
Learn. Time-	Maximizes academic learning time.
Manages-	Effectively manages student behavior.
Logical-	Presents information in a logical sequence.
Cognitive-	Presents information at various cognitive levels.
Clear-	Presents information accurately and clearly.
Strategies-	Utilizes strategies to motivate students and learning.
Environment-	Maintains a supportive environment.

Table 6

Product Moment Correlations of Scores from the COKER, the NTE and Overall Principals' Ratings (N=34)[1]

COKER Variables	CS	GK	PK	Prin.
Focus	-04	-21	00	32*
Organizes Presentations	-08	-08	-14	42**
Directions	-20	-21	-19	-03
Participation	28	23	35*	35*
Evaluates	03	-15	11	39*
Organizes Classroom	-32	-50**	-29	32*
Max. Learn. Time	-40*	-38*	-28	11
Manages	-17	-23	03	33*
Logical	05	-12	13	22
Cognitive Levels	35*	24	31	22
Clear Presentations	11	-07	10	37*
Strategies	-03	10	22	21
Environment	-26	26	15	-06

[1]Decimal points omitted.
* $p < .05$
** $p < .01$

Correlations obtained between scores from the observable performance factors from derived from the COKER and scores from the Core Battery of the

NTE appeared to have little meaning. A correlation of .35 ($p<.05$) was noted between Professional Knowledge and classroom participation. Also the correlation of 0.35 ($p<.05$) between Communication Skills and presenting materials at various cognitive levels may have some meaning. Assuming that the COKER is a reliable and valid instrument for measuring classroom performance the use of the NTE in predicting success in the classroom was questioned.

Several significant correlations ($p<.05$) were noted between principals' overall ratings of performance and the variables of focus of the lesson, organization of presentations, class participation, evaluation, organization of the classroom, classroom management, and presentation of materials. However, Coker, Medley and Soar (1980), have questioned the validity and reliability of principals ratings of teachers performance. Correlations between the F-scale scores and the 13 factors were also examined. The correlations were low and were of no signifiance. This would be expected.

Summary and Conclusions

The longitudinal study of the teacher education programs of Tennessee Technological University entered it 14th year in 1986-87. The activities of the project were combined with the Center for Teacher Education Evaluation, a research and development project designed to develop models for the evaluation of teacher education programs.

The COKER was the primary research instrument utilized for gathering performance data on the graduates of the teacher education program. Also, personal data, and principal ratings of performance were obtained. Due to scoring problems the data derived from the COKER were of little use in program improvement and planning. The data derived from other sources were comparable with that derived from earlier years of the project.

The COKER is an excellent, well developed instrument. The instrument when properly used and under ideal conditions could provide much information that would be useful in improving programs. However, it has many limitations that make the use of the instrument difficult by many institutions. Following is a summary of some of the perceived problems.

1. Cost of use of the instrument are high. Individuals must be highly trained (50 to 60 hours) and the number of qualified trainers is limited.

2. Collection of data is difficult. Repeated trips must be made to classrooms. If subjects are located any distance from an institution, the required resources for travel can be excessive.

3. Scoring of the instrument is difficult. There are only two or three institutions in the country that can score the instrument. The cost is excessive for the amount of information received.

Based on the experiences of the use of the COKER, the decision was made to drop the instrument in favor of data collection systems that would be more

reasonable in terms of needed resources. If the resources are available (personnel, time, travel funds, and patience), the COKER is an excellent instrument.

References

Adorno, T. W., et al. (1950). The authoritarian personality. New York: Harper.

Ayers, J. B. (1986). TennesseeTechnological University Teacher Education Evaluation Model-Year XIII. (Report 86-2). Cookeville, TN: Tennessee Technological University, Center for Teacher Education Evaluation. (ERIC Document Reproduction Service No ED 275 635)

Brown, B. (1970). Experimentation in teaching practice. Journal of Research and Development in Education, 4, 14-22.

Coker, H., & Coker, J. G. (1982). Classroom Observations Keyed for Effectiveness Research-User's Manual. Atlanta: Georgia State University/Carroll County Teacher Corps Project.

Coker, H., Medley, D. M., & Soar, R. S. (1980). How valid are expert opinions about effective teaching? Phi Delta Kappan, 62(2), 131-134.

Dickson, G. E., & Wiersma, W. (1980). Research and evaluation in teacher education: A concern for competent, effective teachers. Toledo, OH: The University of Toledo, The Center for Educational Research and Service.

Medley, D. M. (1973). Observation Schedule and Record Form 5 (Verbal). Charlottesville, VA: University of Virginia.

Medley, D. M. (1977). Teacher competencies and teacher effectiveness: A review process. Washington: American Association of Colleges for Teacher Education.

Medley, D. M., Coker, H., & Soar, R. S. (1984). Measurement-based evaluation of teacher performance. New York: Longman.

Medley, D. M., et al. (1981). Assessing teacher performance from observed competency indicators defined by classroom teachers. Journal of Educational Research, 74, 197-216.

Sandefur, J. T. (1970, September)). An illustrated model for the evaluation of teacher education graduates. Washington: American Association of Colleges for Teacher Education.

Say, M. W., Georgiades, R., Porter, C., & Goldstein, J. (1987, February). Teacher appraisal in the Humble I.S.D. a low inference approach to human resource development. Presented at the meeting of the Association of Teacher Educators, Houston, TX.

Soar, R. L., Soar, R. M., & Rogosta, M. (1971). Florida Climate and Control System (FLACCS): Observers' Manual. Gainesville, FL: Institute for the Department of Human Resources.

Spaulding, R. L. (1976). Spaulding Teacher Activity Rating Schedule (STARS). San Jose, CA: San Jose State University.

Wiersma, W. (1986, April). Teacher evaluation for certification: The use of observation systems. In Wiersma, W. (ed.). Issues in Certification, Advancement and Instruction. Toledo, OH: The University of Toledo, The Center for Educational Research and Services.

APPENDIX

CLASSROOM OBSERVATIONS KEYED for EFFECTIVENESS RESEARCH©

1981

SUBJECT	CODES
01	01 – Music & Art
02	02 – Lang. Arts
03	03 – Literature
04	04 – Math
05	05 – Reading
06	06 – Science
07	07 – Soc. Stu.
08	08 – Planning
09	09 – Transition
10	10 – Other

DISTRICT	SCHOOL	TEACHER	CLASS SIZE	DATE MO.	DATE DAY	DATE YEAR	RECORDER	GRADE	VISIT

SECTION A

	STUDENT	1. Passive compliance	2. Active compliance	3. Answers quest. directed to him	4. Answers quest. voluntarily	5. Substantive question	6. Procedural question	7. Volunteers comment	8. Off-Task Questions or comments	9. Individual Off-Task (V or Non V)	10. Disruptive Off-Task (V or Non V)
PRESENTING	11. Orienting										
	12. Motivates: Prob. struct. (set induction)										
	13. Directs – learning related										
	14. Directs – non learning related										
	15. Explains, discusses, tells										
	16. Digresses										
	17. Praises										
QUESTIONING	18. Open ended No wrong answer										
	19. Recall info information										
	20. Use or Application										
	21. Ampl./Eval./Elaborate – Students' own idea										
	22. Ampl./Eval./Elaborate – Other Students' idea										
	23. Status										
RESPONDING	24. Praises with explanation										
	25. Praises without explanation										
	26. Tells, gives info										
	27. Wrong answer/No answer/Gives info										
	28. Accepts – neutral checks own perception										
	29. Uses, extends										
	30. Waits (long)										
	31. Focusing Cueing										
	32. Asks another student										
	33. Repeats after feedback										
	34. Listens: Observes Pays attention										
	35. Suggests/Questions/Directs/Ignores										
	36. Interrupts Rejects										
	37. Criticizes/Commands										
	38/39 Code Involvement								1 2 3 4 5		1 2 3 4 5

KEY
1 = less ¼ of class
2 = ¼ of class
3 = ½ of class
4 = ¾ of class
5 = all of class

Joan G. & Homer Coker
Georgia State University
College of Education
Atlanta, GA 30303

DO NOT WRITE IN THIS AREA

02062

NCS Trans Optic M10 71269 321 A1301

SECTION B

STUDENT

○	10. Enthusiastic	
○	11. Praises another	
○	12. Pats, hugs Pos. horseplay	
○	13. Laughs, smiles	
○	14. Shows pride	
○	15. Agrees, V or NV	
○	16. Friendly, V or NV	
○	17. Helps Shares/Consider V or NV	
○	less ± of class	48. Code Involvement (10-47)
○	± of class	
○	± of class	
○	± of class	
○	all of class	
○	49. Intense Involvement	
○	50. Leadership	
○	51. Follows Routine WO Rem.	
○	52. S. answers another S.	
○	53. Works w. social	
○	54. Prefers work/play	
○	55. Confuses work/play	
○	56. Confused	
○	57. Self-directed, Inappro.	
○	58. Wanders about	
○	59. Pouts, withdraws	
○	60. Shows fear, shame	
○	61. Makes face, frowns	
○	62. Tattles	
○	63. Teases	
○	64. Resists	
○	65. Picks at another/Neg. horseplay	
○	66. Dem/Com/Boss/Con	
○	67. Dis/Annoy/Interferes	
○	68. Takes/Damage/Stamp/Throws	
○	69. Cut/Disparages	
○	70. Att/Hit/Hurt/ ~ agg	
○	less ± of class	71. Code Involvement (61-70)
○	± of class	
○	± of class	
○	± of class	
○	all of class	
○	less ± of class	72. Code Interest/ Attention
○	± of class	
○	± of class	
○	± of class	
○	all of class	

TEACHER

METHODOLOGY

Motivation	○	73. Intrinsic Immed
	○	74. Intrinsic Future
	○	75. Extrinsic
Student Choice	○	76. S no choice
	○	77. S lim choice
	○	78. S free choice
Supervision	○	79. Aloof, detached
	○	80. Observes, Monitor
	○	81. Joins, partic.
	○	82. Manages simul. Act.
	○	83. Close supery.
Focus	○	84. T. O. Prob.
	○	85. S. O. Prob.
Source	○	86. Text, pkg. res.
	○	87. Multiple res.
Differ/ Evaluation	○	88. Same mat./eval.
	○	89. Indiv. mat./eval.
	○	90. S. partic. eval.
Student Expres.	○	91. Discourages
	○	92. Encourages
Student Perpl.	○	93. Prevents
	○	94. Explains
Mis-inform.	○	95. Accepts
	○	96. Corrects
Strategy	○	97. Inductive
	○	98. Deductive
	○	99. Transductive
	○	100. Expository
Cogn. Level	○	101. Simple
	○	102. Complex
Use of Student time	○	103. T. assigns SA
	○	104. Indep. work

AFFECT — NON-VERBAL — CONTROL

○	105. Warm, cong.		
○	107. Nod, smiles	○	108
○	109. Tou/Pat/Hug	○	110
○	111. Pause	○	112
○	113. Eye contact	○	114
○	115. Ignore	○	116
○	117. Gestures, Sig'l Raps	○	118
○	119. Shakes head/shh	○	120
○	121. Takes something	○	122
○	123. Glares, Frowns	○	124
○	125. Holds, Pushes	○	126

VERBAL

○	127. Praise, non-sub.	○	128
○	129. Agrees, Supports	○	130
○	131. Pos. Indiv. Attn.	○	132
○	133. Reminds	○	134
○	135. Says stop	○	136
○	137. Firm, Sharp	○	138
○	139. Sco/Warns, Pun.	○	140

GROUPING (Blacken ⓓ for dominant mode)

	Non pres.			No. of Stu.	Prescribed			
	Wt.		WOT		WT.		WOT	
141	○ ⓒ	146	○ ⓓ	1	○ ⓦ	150	○ ⓓ	155
142	○ ⓒ	147	○ ⓓ	2-3	○ ⓦ	151	○ ⓓ	156
143	○ ⓒ	148	○ ⓓ	4-[illegible]	○ ⓦ	152	○ ⓓ	157
144	○ ⓒ	149	○ ⓓ	[illegible] +	○ ⓦ	153	○ ⓓ	158
145	○ ⓒ			ALL	○ ⓦ	154	○ ⓓ	159

COMMENTS:

DO NOT WRITE IN THIS AREA

I. Teaching Skill
- 1. establishes focus of the lesson
  - a. secures students' attention at beginning of lesson/activity
  - b. presents objectives
  - c. previews sequence of activities
  - d. reviews previous lesson
  - e. gives/seeks rationale for lesson
  - f. relates lesson to previous or future lesson
- 2. organizes presentations
  - a. actively involved in presenting content
  - b. stresses main points
  - c. demonstrates skills and processes to be learned
  - d. provides opportunities for practice (guided/independent)
  - e. provides closure
- 3. gives clear and easily understood directions
  - a. explains steps to be followed
  - b. identifies potential areas of difficulty
  - c. presents directions in written format, if appropriate
  - d. clarifies directions and explanations, as needed
  - e. provides examples
- 4. provides opportunities for students to actively and successfully participate
  - a. varies activities between teacher focused and student focused
  - b. accepts student contributions and questions
  - c. provides prompts, as needed
  - d. encourages slow or reluctant students, as needed
  - e. varies activities/materials/assignments to meet different needs
- 5. evaluates and provides feedback on student progress
  - a. asks questions, requests demonstration, checks seatwork
  - b. frequently checks student progress
  - c. provides corrective feedback, as needed
  - d. reinforces correct responses
  - e. reteaches, as needed

II. Classroom Management and Organization
- 1. organizes the classroom
  - a. uses procedures and routines which facilitate instruction
  - b. maintains seating arrangement/grouping appropriate for the activity
  - c. has materials and facilities ready for use
  - d. makes information regarding scheduled activities/assignments available to students
- 2. maximizes academic learning time
  - a. begins lesson promptly
  - b. avoids breaks in lesson flow
  - c. keeps students on task
  - d. avoids digressions
  - e. moves through the lesson at an appropriate pace
- 3. effectively manages student behavior
  - a. specifies expectations for student behavior
  - b. monitors class activity
  - c. reinforces appropriate behavior
  - d. stops and redirects inappropriate behavior
  - e. applies rules consistently and fairly

III. Presentation of Subject Matter
- 1. presents information in a logical sequence
  - a. states facts
  - b. identifies concepts and gives example, as needed
  - c. states generalizations or principles and explains relationship to given concepts
  - d. supports generalizations with facts
  - e. utilizes a variety of teaching strategies, i.e., deductive, inductive, expository, transductive
- 2. presents information at various cognitive levels
  - a. knowledge
  - b. comprehension
  - c. application
  - d. analysis
  - e. synthesis
  - f. evaluation
- 3. presents information accurately and clearly
  - a. makes no obvious errors
  - b. avoids vague terms
  - c. responds knowledgeably to student questions
  - d. uses correct syntax
  - e. pronounces words correctly
  - f. uses vocabulary appropriate to students

IV. Learning Environment
- 1. uses strategies to motivate students for learning
  - a. relates content to student interests
  - b. emphasizes the value of the learning activity or content
  - c. reinforces learning efforts of students
  - d. provides options to students in manner of fulfilling assignments
  - e. states challenge/varies levels of concern
- 2. maintains supportive environment
  - a. models courteous behavior
  - b. uses students' names
  - c. interacts positively with all students in the class
  - d. exhibits sense of humor, smiles, or laughs
  - e. positions self near students
  - f. avoids sarcasm or criticism

HUMBLE I.S.D.
TEACHING BEHAVIORS
AS IDENTIFIED BY COKER KEYS

1. Establishes Focus of the Lesson
Teacher is orienting and motivating while students are waiting or actively complying. Teacher uses both teacher and student questions to provide focus. Teacher asks open-ended questions to which students respond. Teacher uses student ideas. There is an absence of procedural and off-task questions and comments on the part of students.

2. Organizes Presentations
Teacher gives learning related directions and explantions, and provides discussions with students. Teacher uses questioning procedures to insure active involvement in understanding content. This may include asking status to actively demonstrate what is to be learned and providing student choices for practice (both guided and independent). The teacher avoids non-learning related activities and digressions. Students ask substantive questions and make on-task comments. There is an absence of interruptions by the students.

3. Gives Clear and Easily Understood Directions
Teacher explains to class what they are going to do; sets performance goals; directs students in preparatory activities before or during a learning task; and seeks substantive status of students. In addition, the teacher does not repeat procedural or substantive instructions. The students are in various modes of on-task behavior including complying and volunteering. There is an absence of student procedural questions or off-task behaviors. There is also an absence of student confusion.

4. Provides Opportunities for Students to Actively and Successfully Participate
Teacher uses a variety of teaching strategies which create high levels of student interest while the students may be working alone or with others. The teacher interacts and responds in several supervisory modes including close supervision and attending to student activities. The teacher also accepts student contributions and corrects misinformation without criticism.

5. Evaluates and Provides Feedback on Student Progress
Teacher uses strategies which provide for interactions with students including student-initiated questions and comments. There may be a variety of responses by the teacher which provide for feedback and evaluation on the part of teacher. There is an absence of passive compliance and a minimal use of control measures on the part of the teacher.

6. Organizes the Classroom
Teacher moves students from passive to active compliance and demonstrates a wide range of presenting methodologies which include orienting the students to the day's task, directing the students to do something relative to the learning task, and redirecting students when they are involved in some non-learning activity. There is an absence of student off-task questions and comments.

7. Maximizes Academic Learning Time
Students demonstrate task involvement in a variety of ways including intense involvement and a high level of interest/attention. There is an absence of digression on the part of the teacher. Student off-task questions and disruptive behavior are minimal.

8. Effectively Manages Student Behavior
Teacher uses supportive classroom management strategies including status questioning and correcting without criticism. There is an absence of negative verbal and non-verbal teacher behaviors, such as glares, pushes, firm/sharp words, and/or scolds.

9. Presents Information in a Logical Sequence
Teacher uses a variety of strategies including deductive, inductive, transductive, and expository modes while students are on-task and attending. There is an absence of student questions and a minimum of teacher responding behaviors.

10. Presents Information at Various Cognitive Levels
Teacher is in various modes of presenting using simple and complex cognitive levels. Teacher uses multiple levels of questioning.

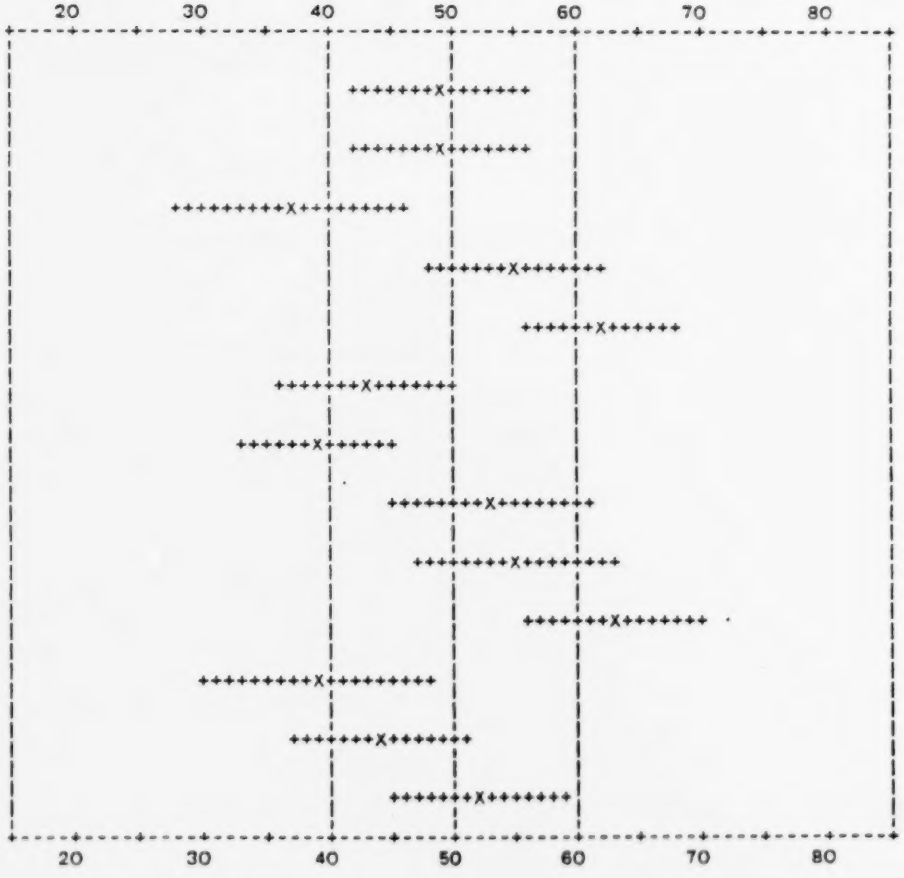
11. Presents Information Accurately and Clearly
Teacher explains, discusses, and tells information to the students while they are in an attending mode. Students may ask substantive questions and the teacher responds by telling or giving information. The teacher also may use the students' comments or expand on a student comment. There is an absence of misinformation and student confusion.

12. Uses Strategies to Motivate Students for Learning
Teacher uses intrinsic and extrinsic motivation techniques and the students are enthusiastic. Student expression is encouraged and student perplexity is fostered. Also, the teacher may join and participate in learning activities as their equal. There is an absence of teacher aloofness as well as a minimal use of control techniques.

13. Maintains Supportive Environment
Teacher is warm and congenial and is willing to nod, smile, be agreeable, support and praise with explanations, and give positive individual attention. Students use praise, laugh, smile, show pride and are friendly and helpful. Teacher listens and observes and gently redirects when the student is off-task. There is an absence of negative teacher behaviors including interrupts, rejects, criticizes or commands as well as student negative behaviors such as pouts, shows fear, resists teacher, or criticizes.

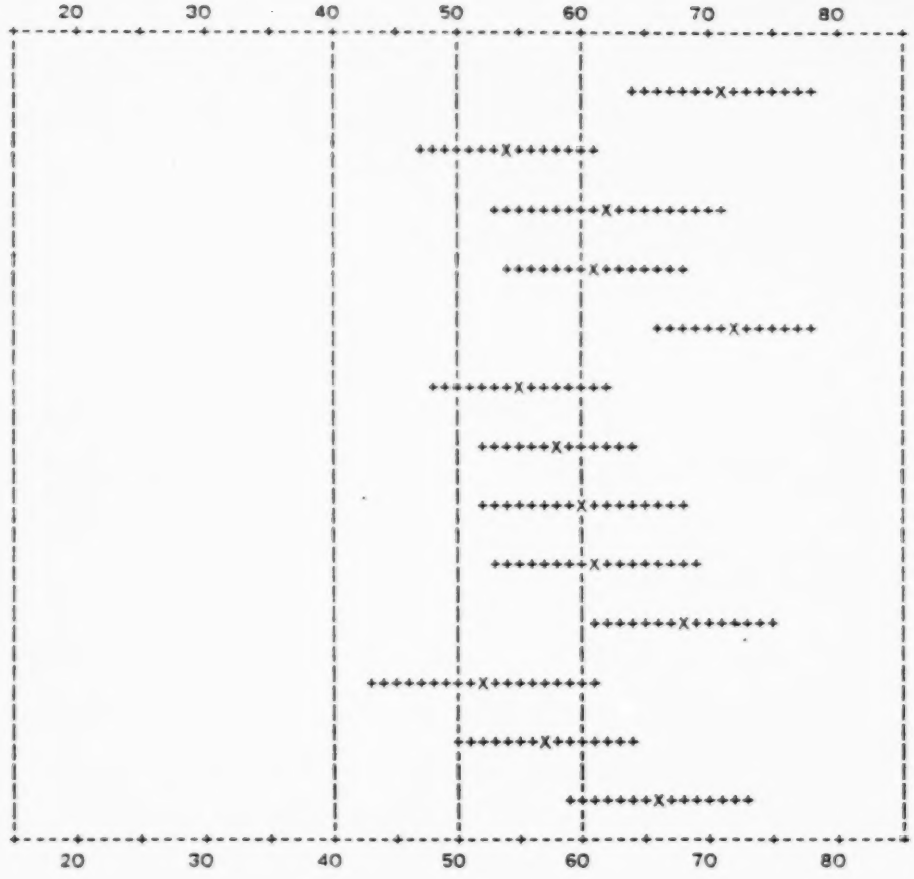
T E A C H E R E F F E C T I V E N E S S P R O F I L E

		SCORE
	TEACHER NUMBER 3350	
1	ESTABLISHES FOCUS OF THE LESSON (1)	49.06
2	ORGANIZES PRESENTATIONS (2)	48.56
3	GIVES CLEAR AND EASILY UNDERSTOOD DIRECTIONS (3)	37.07
4	PROVIDES OPPORTUNITIES FOR STUDENTS TO ACTIVELY AND SUCCESSFULLY PARTICIPATE(4)	54.61
5	EVALUATES AND PROVIDES FEEDBACK ON STUDENT PROGRESS (5)	62.22
6	ORGANIZES THE CLASSROOM (6)	42.96
7	MAXIMIZES ACADEMIC LEARNING TIME (7)	38.79
8	EFFECTIVELY MANAGES STUDENT BEHAVIOR (8) MIDDLE SCHOOL	52.89
9	PRESENTS INFORMATION IN A LOGIGAL SEQUENCE (9) MIDDLE SCH	55.06
10	PRESENTS INFORMATION AT VARIOUS COGNITIVE LEVELS (10)	62.54
11	PRESENTS INFORMATION ACCURATELY AND CLEARLY (11)	39.03
12	USES STRATEGIES TO MOTIVATE STUDENTS FOR LEARNING (12)	43.69
13	MAINTAINS SUPPORTIVE ENVIRONMEN (13)	51.65



T E A C H E R E F F E C T I V E N E S S P R O F I L E

	SCORE
TEACHER NUMBER 3360	
1 ESTABLISHES FOCUS OF THE LESSON (1)	70.96
2 ORGANIZES PRESENTATIONS (2)	54.14
3 GIVES CLEAR AND EASILY UNDERSTOOD DIRECTIONS (3)	61.97
4 PROVIDES OPPORTUNITIES FOR STUDENTS TO ACTIVELY AND SUCCESSFULLY PARTICIPATE(4)	61.11
5 EVALUATES AND PROVIDES FEEDBACK ON STUDENT PROGRESS (5)	72.14
6 ORGANIZES THE CLASSROOM (6)	54.71
7 MAXIMIZES ACADEMIC LEARNING TIME (7)	58.36
8 EFFECTIVELY MANAGES STUDENT BEHAVIOR (8) MIDDLE SCHOOL	59.62
9 PRESENTS INFORMATION IN A LOGIGAL SEQUENCE (9) MIDDLE SCH	61.16
10 PRESENTS INFORMATION AT VARIOUS COGNITIVE LEVELS (10)	67.87
11 PRESENTS INFORMATION ACCURATELY AND CLEARLY (11)	51.88
12 USES STRATEGIES TO MOTIVATE STUDENTS FOR LEARNING (12)	56.70
13 MAINTAINS SUPPORTIVE ENVIRONMEN (13)	65.99



Chapter II

ASSESSMENT OF TEACHING PERFORMANCE AND CLASSROOM OBSERVATION 1988

Jerry B. Ayers

Accountability in education has been the subject of many investigations in recent years. The Center for Teacher Education Evaluation (CTEE) at Tennessee Technological University was established for the purpose of assisting institutions and agencies to improve programs for the preparation of teachers. As an integral part of its activities the CTEE has undertaken selected studies that were designed to help institutions collect valid information for program improvement in more reliable and economical ways.

Evaluation of the product of a teacher education program (i.e., the graduates) is a time consuming and costly process. It is also one of the most important. The most commonly employed method is to observe the graduates in the classroom using one of the standardized observation instruments that have been developed in recent years. It is particularly difficult for small institutions to conduct large scale followup studies of their graduates if direct observation work is involved.

What are some alternatives to gathering direct information from the classroom of the graduates? Students are in the classroom for six to seven hours a day and are the consumers of the work of the teachers. Therefore, this ready made source of information may provide an attractive alternative method (from an economic standpoint) to collect information about graduates of a teacher education program. Also, information collected in this manner can be combined with other data to yield a more accurate picture of teachers and their ability in the classroom. This in turn is useful information for program improvement.

The literature on student assessment of teaching is extensive. Gleason (1986), Overall and Marsh (1982), Dowell and Neal (1982) and Miles (1988) have conducted reviews of the literature on the subject. The results are conflicting and authors indicate there is a need for additional studies in the area (Blackburn & Clark, 1975; Murray, 1983; Payne & Manning, 1987).

What are some of the concerns and problems relative to the use of student evaluations of teaching? Rodin (1973) questioned the reliability of student evaluations of teaching. His argument was based on criticism that the most frequently used methods of determining reliability of student ratings actually can neither prove nor disprove student reliability. Successive evaluations taken close together may introduce a memory factor. Evaluations which were taken over a period of time, could reflect actual differences in teaching yet be interpreted as unreliability of the students as evaluators. However, Overall and Marsh (1982), have studied the problems of reliability of student estimates. They have found reliability coefficents as high as 0.80 when using comparisons of different items (on the same instrument) which measured the same characteristics.

The validity of the use of student ratings has also been questioned. Proponents contended that students are the best judges of a teacher as they are the ultimate consumers (Webb & Nolan, 1955). Wagner (1983) pointed out that students are uninformed of pedagogical practice, and are, therefore, incapable of making sound judgements.

Conclusions about the use of student evaluations of teaching have been as diverse as the methods employed. In recent years, research has supported the conclusion that student ratings are valid and reliable methods for evaluating teachers (Gleason, 1986). The underlying reason for widespread disagreement and inability to compare different studies seems to be that teaching is a hypothetical construct for which there is no general consensus concerning indicators of quality (Marsh, 1984).

What is the relationship between student ratings of teacher and the teachers overt behavior in the classroom? Payne and Manning (1986) examined the relationships between elementary student teachers' verbal behavior and pupil assessment of teaching performance. Instrumentation for the study included the Student Perceptions Scale of the Teacher Performance Assessment Instruments, as a measure of student ratings of teacher behavior, and Flanders Interaction Analysis as a measure of teacher performance in the classroom. They concluded that student ratings should be used in conjunction with direct observation in assessing the performance of teachers. Miles (1988) studied the relationships between selected observable characteristics of secondary teachers and student assessments of teaching performance. The Student Evaluation of Teaching-I (Veldman, 1970) and the Flanders System of Interaction Analysis were the two major data gathering instruments. Miles (1988) concluded that there was some relationship between student ratings and performance as measured by use of the Flanders System. However, the instruments should be used in tandem rather than student evaluations substituting for direct observation. In a study of teachers in grades K-4 utilizing similar instrumentation, Odle (1988) found similar results.

Purpose

The purpose of the present study was to investigate the relationships that existed between student evaluations of teachers' performance and measures of teaching behavior recorded through observation by independent observers. More specifically the study was designed to study graduates of Tennessee Technological University who completed a bachelor's or master's degree in fiscal year 1987 and were teaching in a school in middle Tennessee in the Spring of 1988.

Methodology

The College of Education at Tennessee Technological University initiated in the fall of 1973 an intensive longitudinal follow-up study of the graduates of its teacher education programs. Data have been gathered each year on a sample of graduates who were teaching within a 75 mile radius of the University. During 1987-88, 42 individuals who received the bachelor's or master's degree in fiscal year 1987 volunteered to participate in the study. Twenty-three were teaching above the fourth grade level.

Two instruments were used to gather data in this phase of the study. A modification of the Kentucky Career Ladder Classroom Observation Instrument (KCOI) developed at Western Kentucky University (WKU, 1987) was used by specially trained graduate research assistants to collect data in the classrooms of the subjects. The low inference instrument was developed based on observable teacher behaviors drawn from the literature on teaching. The instrument is divided into two parts. The first requires the observer to tally the frequency at which the teacher is being observed demonstrating the behaviors on the instrument. Each time a behavior which is on the instrument is observed, a tally mark is made. For example, if the teacher asks academic questions which are directly related to the content of the lesson, then each separate question is tallied on the appropriate line. The second part of the instrument requires the observer to respond to a series of statements about the lesson, i.e., was the teacher involved. A trained research assistant visited in the classroom of each subject on three to four occasions to collect data.

Following is a summary of the dimensions measured by the instrument. Questioning can have a score range from -1.00 to +1.00, while all other dimensions have a range from 0.00 to +1.00.

Questions-Asking academic questions has been found to be positively related to student achievement. The range of scores is between -1.00 and +1.00. A score of -1.00 indicated that none of the questions asked were related to the lesson while a score of +1.00 indicated all of the questions asked were related to the lesson.

Call-Outs-A score of 0.00 indicated there were no call-outs in the classroom. Call-outs minimize disruption of the classroom and the lesson.

Responses-This category indicated the degree of response that a teacher made to a student's answer. A score of 1.00 indicated that all of the responses of the teacher were designed to rephrase/amplify or correct/clarify students' comments.

Praise-A score of 1.00 indicated that all of the praise statements made to students during a class were either specific to the performance of an individual or were positive statements made to the group regarding their academic performance.

Communication-A score of 1.00 indicated that all responses made by the teacher to students' answers either gave direction, focused the students' attention on a particular idea or problem or challenged the students to try harder.

Management-Management focused on how teachers handled disruptions in the classroom. A score of 1.00 is obtained by stopping misconduct positively, using conduct related praise or having no disruptions. Stopping misconduct negatively and not stopping misconduct at all hinder the learning process.

Time Management-Time management referred to teacher behavior which demonstrated concern for both teacher and student time-on-task. A score of 1.00 indicated a teacher had minimized management time.

Seatwork-A score of 1.00 was obtained when the teacher moved around the room monitoring student behavior and provided individual explanations to students each time seatwork was assigned. Inappropriate seatwork activity for the teacher included remaining distant from the students, sitting at his/her desk, or to leave the room.

Begins Promptly-A score of 1.00 was obtained when the teacher began instruction promptly in each lesson observed.

Academic Focus-This represented those things that a teacher did to establish a focus for instruction. A teacher obtained a 1.00 for academic focus by reviewing, providing an overview of the lesson content or objectives, and providing an overview of the instructional process to be used.

Presents New Content-This category represented those teacher behaviors associated with the presentation of new content or skill lessons that have been found to be related to positive student outcomes.

Effective Questioning Techniques-A score of 1.00 indicated the teacher demonstrated only positive questioning techniques; i.e., paused before calling on a student, allowed student time to respond, guided the student if needed, and elicited responses from a variety of students in a nonrepetitive pattern.

Provided Clear Directions-A score of 1.00 indicated that each time it was appropriate the teacher provided clear directions for seatwork and/or homework.

Was Involved-Teacher involvement included the amount of time the teacher spent on task. A score of 1.00 indicated the teacher was involved in each lesson that was observed.

Was Enthusiastic-Teacher enthusiasm is crucial to the learning process. A score of 1.00 indicated the teacher demonstrated enthusiasm in each of the lessons that were observed.

The students in one classroom of each subject were asked to complete the Student Evaluation of Teaching-I (SET) developed by Veldman and Peck (Veldman, 1970). The instrument consisted of ten statements about their teacher. The students respond "very much true," "true," 'false," or "very much false." Five dimensions of teaching were measured including: Friendly and Cheerful (SET1), Knowledge and Poise (SET2), Lively and Interested (SET3), Firm Control (SET4), and Non-directive (SET5). In addition a total score (SETT) was computed.

Appropriate descriptive statistics and correlations were computed for the teacher and student observed behaviors. No attempt was made to compare subjects by such variables as degree, sex, or length of teaching experience.

Table 1

CORRELATION MATRIX FOR SELECTED VARIABLES (N=23)*

	$\bar{X}$	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1. SET-1	331	46	100	54	84	-15	69	44	-08	36	-02	25	-06	26	13	22	03	11	36	-05	42	17	33
2. SET-2	345	29		100	43	14	21	24	01	11	09	-11	-19	15	-25	05	-04	04	28	-04	31	23	15
3. SET-3	283	47			100	-31	70	47	-12	21	-29	22	-22	22	17	12	-20	-19	23	-07	21	-02	29
4. SET-4	296	26				100	-10	-13	20	-11	46	-06	-18	-27	-08	-36	38	47	18	04	33	29	-06
5. SET-5	254	51					100	44	00	22	-12	-02	-16	36	25	-06	04	-12	10	-19	25	09	19
6. SET-T	308	86						100	07	10	02	-11	-19	14	12	08	06	27	23	16	16	-19	29
7. QUESTIONS	1.70	0.35							100	-15	28	11	-04	16	-23	-04	-01	-07	23	38	-03	23	08
8. CALL-OUTS	0.18	0.15								100	26	07	-10	-23	07	-11	-01	06	-04	-50	35	04	14
9. RESPONSES	0.94	0.22									100	-07	-09	-05	-05	15	07	45	26	26	47	21	-07
10. PRAISE	0.22	0.27										100	05	01	-06	-10	-14	12	14	09	14	-24	16
11. COMMUNICATION	0.98	0.05											100	17	-17	39	-06	04	03	10	-15	13	20
12. MANAGEMENT	0.90	0.20												100	29	15	10	-06	20	41	-17	-04	25
13. TIME MANAGEMENT	0.75	0.33													100	-13	24	21	-07	-12	-04	-03	17
14. SEATWORK	0.75	0.41														100	-10	03	02	31	21	-04	05
15. BEGINS PROMPTLY	0.84	0.27															100	27	30	14	-02	16	19
16. ACADEMIC FOCUS	0.50	0.23																100	29	17	45	14	26
17. PRESENTS NEW CONTENT	0.75	0.25																	100	52	21	44	29
18. EFFECTIVE QUESTIONING	0.74	0.26																		100	-02	16	31
19. PROVIDED CLEAR DIRECTIONS	0.86	0.30																			100	16	-05
20. WAS INVOLVED	0.93	0.17																				100	24
21. WAS ENTHUSIASTIC	0.37	0.41																					100

*Decimal points on all correlations have been omitted. Underline indicates a correlation significant at or beyond the .05 level.

Results

The results of the study were combined and shown in Table 1. The mean and standard deviations for the administration of the SET were similar to those previously reported over the past 15 years (Ayers, 1986). Therefore, it was felt that the student reactions toward the subjects in the study were probably reliable and valid. An examination of the means and standard deviations of the teaching behaviors of the subjects indicated they were comparable to those gathered using the instrument in the State of Kentucky (Craig, 1988).

The product-moment correlations were also shown in Table 1. Correlations between the various factors of the SET were similar to those from earlier studies of the use of the instrument (Ayers, 1986). Correlations among the various factors of the KCOI were as would be expected.

Significant correlations between scores from the Friendly and Cheerful (SET1) factor of the SET and the KCOI were confined to three dimensions. There were low but significant correlations with the teacher observed dimensions of call-outs, presentation of new material, and providing of directions. No explanation can be offered for these results.

Scores from the SET4 dimension of Firm Control correlated significantly and positively with the teacher observed factors of responses and begins class promptly. A significant negative correlation was found with seatwork. These correlations would stand to reason. If a teacher demonstrated more control of the class, one would expect the class to begin promptly, respones from the teacher would be to the point, and the teacher would control the class through circulation during seatwork and similar activities.

Scores from SET5 dimension of Non-directive correlated significantly with the teacher dimension of management. This indicated that the observers and students were perceiving the teacher in the same manner relative to management style in the classroom.

Conclusions

Oliver and Taylor (1980) pointed out that the connections between teacher characteristics and teaching processes were increasing. In the present study teacher characteristics were delineated by use of the Kentucky Career Ladder Classroom Observation Instrument. Student reactions toward the teacher were obtained by use of the Student Evaluation of Teaching-I. The results of the study indicated that there were few relationships of significance between scores from the two instruments. Assuming the two instruments used in the study were valid and reliable, one would have expected more significant correlations between student and independent observer judgements of the subjects in the study.

The findings of the present investigation corroborate other studies of a similar nature (Miles, 1988; Odle, 1988; Payne & Manning, 1986). In evaluating teaching, there is a need for a variety of instruments and techniques. In spite of the fact the students are in the classroom for up to

seven hours a day, they may still not be able to judge fully the quality of the teacher.

The present study was conducted using a small set of subjects who had volunteered to participate in a larger study. Therefore, there could well be a sample bias. On the surface the two instruments would appear to measure similar characteristics. A detailed factor analysis may, however, reveal that the instruments are in fact measuring several different traits of a teacher. Replication of this study with larger groups of teachers and in more controlled settings is needed.

The study does point out that use of multiple measures to evaluate teachers is still needed. Additional work may lead to the development of a system that could aid in the evaluation of teachers in a manner that is more cost effective and would thus allow even small institutions to have information about their graduates that would be useful for program improvement.

References

Ayers, J. B. (1986). Tennessee Technological University Teacher Education Evaluation Model-Year XIII. (Report 86-2). Cookeville, TN: Tennessee Technological University, Center for Teacher Education Evaluation. (ERIC Document Reproduction Service No. ED 275 635)

Blackburn, R. T., & Clark, M. J. (1975 Spring). An assessment of faculty performance: Some correlates between administrator, colleague, student and self-ratings. Sociology of Education, 48, 242-256.

Craig, J. R. (1988). [Pilot test of the Kentucky Career Classroom Observation Instrument]. Unpublished descriptive statistics.

Dowell, D. A., & Neal, J. A. (1982). A selective review of the validity of student ratings of teaching. Journal of Higher Education, 53(1), 51-61.

Gleason, M. (1986). Getting a perspective on student evaluation. AAHE Bulletin, 39(1), 10-13. (ERIC Document Reproduction Service No. ED 265 801)

Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. Journal of Educational Psychology, 76(5), 707-754.

Miles, S. K. D. (1988). A study of relationships between selected secondary teacher variables and student assessments of teaching performance. Unpublished master's thesis. Tennessee Technological University, Cookeville, TN.

Murray, H. G. (1983). Low-inference classroom teaching behaviors and student ratings of college teaching effectiveness. Journal of Educational Psychology, 75(1), 138-149.

Odle, A. (1988 May). Relationships between elementary teachers' behavior and student assessment of teaching performance. Unpublished master's problem paper. Tennessee Technological University, Cookeville, TN.

Oliver, B., & Taylor, J. L. (1980). Teacher characteristics and classroom behavior. The Journal of Classroom Interaction, 14(1), 11-18.

Overall, J. U., & Marsh, H. W. (1982). Students' evaluations of teaching: An update. AAHE Bulletin, 35(4), 9-13. (ERIC Document Reproduction Service No. ED 225 473)

Payne, B. D., & Manning, B. H. (1986). Relationships between elementary student teachers; verbal behavior and pupil assessment of teaching performance. Journal of Classroom Interaction, 21(2), 26-30.

Rodin, M. J. (1973). Research: Can students evaluate good teaching? Change, 5(6), 66-67.

Veldman, D. J. (1970). Student Evaluation of Teaching (Research Methodology Monograph No. 10). Austin, TX: R&D Center for Teacher Education.

Wagner, P. A. (1983). Classes of legitimate evidence for identifying effective teaching. Clear Lake City: University of Houston. (ERIC Document Reproduction Service No. ED 227 735)

Webb, W. B., & Nolan, C. Y. (1955). Student, supervisor, and self-ratings of instructional proficiency. Journal of Educational Psychology, 46(1), 42-46.

Western Kentucky University (1987). Manual for observing teacher behavior with the Kentucky Career Ladder Classroom Observation Instrument. Bowling Green, KY: Western Kentucky University.

Chapter III

ASSESSMENT OF TEACHING PERFORMANCE AND CLASSROOM OBSERVATION 1989

Jerry B. Ayers

Longitudinal follow-up studies are important to determine the effectiveness of teacher education preparation programs. In 1973, Tennessee Technological University initiated an intensive longitudinal follow-up study of the graduates of its teacher education programs. Data have been gathered each year on a sample of graduates who were teaching within a 75 mile radius of the University. During 1988-89, the project was continued for a 16th year.

The Center for Teacher Education Evaluation was under staffed in 1988-89 for the number of commitments that had been undertaken. Therefore, it was impossible to start the follow-up project and actual classroom visitation until the Spring of 1989. As a result teachers were reluctant to volunteer to participate in the project since it was near the close of the school year. Fifteen individuals who received the bachelor's or master's degree in fiscal year 1988 did agree to participate in the study. Although the number of participants was small it was felt that useful data could be collected.

Chapters I and II of this report detailed the use of the Classroom Observation Keyed for Effectiveness Research (COKER) as the primary research tool in 1986-87. In 1987-88, (Chapter II), the Kentucky Observation Instrument (KOI) was utilized as the instrument for classroom observation. Based on the two trial years, the KOI was chosen to be used during 1988-89. The instrumentation and methodology were identical to that reported in Chapter II. This information has been omitted, in order to conserve space.

Results

The correlation pattern for the complete set of data derived from the various instruments used in the study were similar to those reported in Chapter II. These data have been omitted from this Chapter. The remainder of this Chapter is divided into four sections (Personal Data, Principal Evaluations, Student Evaluations and Observations in the Classroom) in order to make comparisons of data across time and between the various groups.

Personal Data

Personal data were collected on the graduates from University records and from the administration of the California F-Scale Forms 45 and 40 [see Chapter VI for a detailed study of the use of the F-Scale at Tennessee Technological University]. In order to simplify the findings, only comparisons were made with the 1987-88 B.S. graduates.

The mean grade point for the 1987-88 B.S. graduates who participated in the study was 3.15 (SD=0.39, N=20). In comparison the mean for the 1988-89 graduates was 3.12 (SD=0.40, N=14). These mean grade point averages were comparable with those of other groups who have participated in the follow-up studies conducted in earlier years. The mean F-scale score for 14 1987-88

graduates was 104 (SD=16), while the mean score for the 1988-89 graduates was 107 (N=14, SD=16). These scores were comparable with those reported in earlier years of the study.

Table 1 shows a comparison of National Teacher Examinations (NTE) scores and scores derived from the Pre-Professional Skills Tests (PPST) [See, Chapter V provides a summary of a detailed study of the relationships between the NTE, the PPST and other admission and outcomes assessment measures related to the teacher education programs of the University]. Inferential statistics have been omitted because of the small sample sizes.

Table 1

Comparison of Standardized Test Scores for 1988 and 1989 Follow-up Study Participants

Test	1987 $\bar{X}$	1987 SD	1988 $\bar{X}$	1988 SD
Pre-Professional Skills Tests (1987, N=7, 1988, N=5)				
Reading	174	4.5	181	6.6
Writing	175	6.0	177	1.4
Mathematics	176	3.3	180	4.1
National Teacher Examinations (1987, N=18, 1988, N=5)				
Communication Skills	660	8.4	661	11.9
General Knowledge	657	9.9	659	14.7
Professional Knowledge	666	10.7	659	10.7

Principal Evaluations

A four item questionnaire was completed by the principals of the subjects in the study. Table 2 shows a comparison of the mean ratings for the two groups of bachelor's level graduates. In general the rating were comparable with those reported in earlier years of the study.

Student Evaluations

Student evaluations of teachers provide another dimension for gathering useful information for program improvement. Two forms of an instrument have been used in past years to gather data. One instrument is designed to be used in grades K through 3, while the second instrument is designed for use above the third grade. Because of limited sample size only the data gathered with classes above grade three, in which the Student Evaluation of Teacher was administered are reported in Table 3. Scores tended in both groups to be

somewhat below those reported in earlier years of the use of the instrument. These score differences may be due in part to the small sample sizes involved.

Table 2

Comparison of Principal's Ratings

Area	1988 (N=19) $\bar{X}$	1988 (N=19) SD	1989 (N=10) $\bar{X}$	1989 (N=10) SD
Subject Matter Competence	4.1	0.6	4.1	0.7
Relations With Students	3.9	1.1	4.2	0.9
Appropriateness of Assignments	4.0	0.7	4.2	0.8
Overall Effectiveness	3.9	0.7	4.0	0.8

Table 3

Comparison of Scores Derived from Administration of the Student Evaluation of Teacher

Factor	1988 (N=11) $\bar{X}$	1988 (N=11) SD	1989 (N=7) $\bar{X}$	1989 (N=7) SD
Friendly and Cheerful	317	55	316	34
Knowledgeable and Poised	332	25	338	23
Lively and Interesting	276	55	278	28
Firm Control (Discipline)	294	28	286	40
Non-Directive (Democratic Procedure)	250	55	213	41
Composite Score	293	35	286	10

Classroom Observation

Table 4 shows a comparison between the means and standard deviations for each of the major variables from the KOI for those subjects that received their degrees in 1987-88 and 1988-89. Also shown is the mean scores achieved by a sample of 337 teachers in Kentucky who participated in the initial testing of the instrument. No attempt has been made at this time to make detailed comparisons (inferential statistics) between the groups. Only gross comparisons were made across the two time periods and with the data derived from teachers studies in Kentucky.

The score values shown in Table 4 are relative and can range from 0 to +1.0 The closer a score approaches +1.0, for a particular behavior, the more the behavior or skill is being demonstrated by a teacher. In general, the participants in comparison with the Kentucky teachers were demonstrating the measured behaviors (skills) in a favorable manner. It appeared that both

groups of graduates were rated lower than the Kentucky teachers in the areas of use of praise, management of class time, and in the use of and conducting seatwork. These might be areas of concerns for future methods of teaching classes.

Table 4

Comparison of Means and Standard Deviations for Variables from the KOI for 1988 and 1989 with Data Gathered in Kentucky Field Trails of Instrument

KOI Variable	Kentucky		1988 (N=23)		1989 (N=14)	
	Mean	SD	Mean	SD	Mean	SD
Questions	0.61	0.40	0.69	0.35	0.69	0.14
Responses	0.92	0.28	0.94	0.22	0.90	0.27
Praise	0.34	0.35	0.22	0.27	0.25	0.26
Communication	0.89	0.30	0.98	0.05	0.94	0.10
Management	0.90	0.23	0.90	0.20	0.85	0.20
Time Management	0.86	0.31	0.75	0.33	0.78	0.37
Seatwork	0.93	0.24	0.75	0.41	0.86	0.31
Begins Promptly	0.90	0.13	0.84	0.27	0.97	0.28
Academic Focus	0.66	0.15	0.50	0.23	0.59	0.20
Presents New Content	0.83	0.12	0.75	0.25	0.82	0.18
Effective Questioning	0.80	0.18	0.74	0.26	0.77	0.26
Provided Clear Dir.	0.88	0.16	0.86	0.30	0.88	0.28
Was Involved	0.96	0.09	0.93	0.17	0.94	0.16
Was Enthusiastic	0.40	0.27	0.37	0.41	0.62	0.41

Summary and Conclusions

The application of the basic follow-up model provided information that can be used for program planning and improvement. Although the data collected were somewhat incomplete, it was demonstrated that the system will provide useful information for improving the programs of the University. The basic plans that were further pilot tested during 1987-88 and 1988-89, will be implemented on a full scale basis during each of the next five years. Changes in the basic design will be made only when proven instruments are introduced that will provide more useful information.

Chapter IV

THE NTE, PPST AND CLASSROOM PERFORMANCE

Jerry B. Ayers

The National Teacher Examinations (NTE) were first introduced in 1940 as a measure of an individual's potential performance as a teacher in the classroom. Over the years these tests have been revised and have come into wide spread use Rudner (1988) reported that 22 states required prospective teachers to complete the NTE prior to licensure and additional states were considering adoption of the requirement. In 1983 a companion test, the Pre-Professional Skills Tests (PPST) was introduced as a measure of a student's ability in the basic areas of reading, writing, and mathematics. The PPST is being used by 13 states as one criteria for admission to formal study in professional teacher education programs (Rudner, 1988).

In the past five years there has been a renewed interest in the preparation of teachers and numerous reforms are occurring in teacher education programs. Legislators and others are seeking ways to be accountable to the public and to "improve the quality of teachers in the classrooms." As a result a interest has been shown in utilizing test scores to justify the quality of higher education. The PPST is becoming more widely used as a requirement for admission to study in professional teacher education programs and prospective teachers are being required to achieve higher and higher scores on the NTE as a requirement to achieve a license or certificate to teach. In effect achievement on the PPST has become a requirement for licensure as a teacher.

Many individuals have questioned the use of the PPST and NTE as absolute measures of an individual's ability as a classroom teacher. The true test of an individual as a teacher is the ability to perform in the classroom. Direct observation in the classroom is the key to determining performance.

Ayers and Qualls (1979) examined the concurrent and predictive validity of the National Teacher Examinations(NTE) that were in use during the period from 1974 through 1978. More specifically they examined the relationships between scores on the NTE achieved by a sample of graduates of the teacher education programs of Tennessee Technological University, who were teaching in their first year and, respectively, ratings made by their principals, student ratings of the graduates, and ratings made by independent observers utilizing object scored instruments. The observation instruments were specifically designed to measure the classroom climate and the interaction between pupils and teacher. The instruments included Flanders Interaction Analysis (Amidon & Flanders, 1971), the Classroom Observation Record (Ryans, 1960), and the Tuckman Teacher Feedback Form (Tuckman, 1976).

In 1984 and 1985, Ayers (1988) replicated the earlier study utilizing the same instrumentation for observation and ratings of performance in the classroom and a sample of graduates of Tennessee Technological University who had completed the revised version of the NTE. The results from both the 1979 and 1988 studies were quite similar. The relationships between scores from

the NTE and, respectively, rating devices and observation instruments were low and not significant. If it was assumed that the data collected through the use of the observation instruments and the other devices was valid, the NTE was of little value in predicting success in the classroom.

Siebert (1989) and Ayers (1988) examined the relationships between the NTE and the PPST. As would be expected there were moderate to strong relationships between the three tests of the NTE Core Battery and the three sub-test of the PPST. Correlations were on the order to 0.5 to 0.8. The PPST appeared to be a reasonable predictor of scores on the NTE.

The Center for Teacher Education Evaluation at Tennessee Technological University has been making a series of studies of various types of observation instruments that can be used to determine the level of performance of teachers in the classroom. The observation instruments are based on established principals of assessment and appear to provide a measure of an individuals ability to perform in the classroom. During the past two years the Center has collected observation data on graduates of both bachelors' and master's degree programs utilizing two low inference observation instruments. The Classroom Observation Keyed for Effectiveness Research instrument (COKER) developed by Coker and Coker (1981) and a modification of a low inference observation instrument utilized in the Kentucky Career Ladder program (KOI) developed at Western Kentucky University (1987) have formed the basis for the research.

Purpose of Study

The purpose of the current study was to continue the work initiated in 1978 to study the relationship between scores from the NTE and scores achieved from the use of objective scored observation instruments. A secondary purpose of the study was to examine the relationship between scores achieved by a sample of graduates on the PPST and the results of the use of observation instruments.

Method

During 1986-87 and 1987-88, 100 graduates of the teacher preparation programs of Tennessee Technological University participated in the studies. The sample included 60 who had completed the bachelor's degree and were in their first year of teaching and 40 who had completed the Master of Arts degree and were in the classroom during the first year after completing the degree. Only those individuals who had completed the bachelor's degree served as subjects for this study. Thirty-four subjects from the Spring of 1987 and 28 from the Spring of 1988 had completed the Core Battery of the NTE. Also, 27 subjects in the 1988 group had completed the PPST as a part of the process of admission to the professional teacher education programs of the University.

A specially trained graduate research assistant visited in the classroom of each graduate on three to five occasions during the Spring of the year. The COKER was used in the Spring of 1987 while the KOI was used in 1988. Scoring of the instruments allowed for the determination on an absolute scale of observational teaching behaviors. In addition, such measures as grade point averages, scores from a personality test, and children's ratings of

performance were obtained on each subject in the study. Pearson Product Moment Correlations were computed between classroom performance measures and scores derived from the NTE, the PPST, and the other measures noted above.

Following is a summary of the symbols and corresponding variables drawn from, respectively, the COKER (Table 1) and the KOI (Table 2). It should be noted that the two instruments measure similar traits that a teacher should exhibit in the classroom.

Table 1

Summary of Characteristics Measured by the COKER

Focus-	Establishes focus of the lesson.
Org. Pres.-	Organizes classroom presentations.
Directions-	Gives clear and easily understood directions.
Participation-	Provides opportunities for students to actively and successfully participates.
Evaluates-	Evaluates student progress and provides feedback.
Org. Class.-	Provides organization in the classroom.
Learn. Time-	Maximizes academic learning time.
Manages-	Effectively manages student behavior.
Logical-	Presents information in a logical sequence.
Cognitive-	Presents information at various cognitive levels.
Clear-	Presents information accurately and clearly.
Strategies-	Utilizes strategies to motivate students and learning.
Environment-	Maintains a supportive environment.

Table 2

Summary of Characteristics Measured by the KOI

Questions-	Utilizes academic questions in an appropriate manner.
Responses-	Degree of responses made to student answers.
Praise-	The degree and amount of praise used by the teacher.
Communications-	Thedegree that responsesbyateacher to students' gave direction, focused their attention orchallenged themto try harder.
Management-	Use of conduct related praise.
Time Manage.-	Concern the teacher had for time-on-task.
Seatwork-	Teacher used seatwork in an effective manner.
BeginsProm.-	The degree towhichtheteacher began classwork promptly.
Focus-	Establishment of academic focus in the classroom.
New Content-	Demonstration of skills related to the presentation of new content.
Questioning-	Demonstration of the use of effective questioning techniques.
Directions-	Provided clear directions during teaching.

Table 4 Cont.

Involved-	Demonstration of involvement in the lesson.
Enthusiastic-	The degree to which the teacher demonstrated enthusiasm for learninginthe classroom.

Results

The results of the study were divided into two sections. The first summarized the results obtained in the studies utilizing the COKER. The second section described the results obtained by use of the KOI.

COKER Results

Data were available from a sample of 34 individuals who had completed the bachelor's degree and were in their first year of teaching. Mean scores achieved on the Core Battery of the NTE were as follows: Communication Skills (CS) 663, SD=11; General Knowledge (GK) 659, SD=11; and Professional Knowledge (PK) 661, SD=12. The overall rating of performance of each subject by the building principal was also used in this study. The mean score on a five point scale (1=low to 5=high) was 4.4 (SD=0.7). Table 3 shows a summary of the correlations between the 13 factors from the COKER and the NTE Core Battery scores and the overall principals rating.

Table 3

Product Moment Correlations of Scores from the COKER, the NTE and Overall Principals' Ratings (N=34)[1]

COKER Variables	CS	GK	PK	Prin.
Focus	-04	-21	00	32*
Organizes Presentations	-08	-08	-14	42**
Directions	-20	-21	-19	-03
Participation	28	23	35*	35*
Evaluates	03	-15	11	39*
Organizes Classroom	-32	-50**	-29	32*
Max. Learn. Time	-40*	-38*	-28	11
Manages	-17	-23	03	33*
Logical	05	-12	13	22
Cognitive Levels	35*	24	31	22
Clear Presentations	11	-07	10	37*
Strategies	-03	10	22	21
Environment	-26	26	15	-06

[1]Decimal points omitted.
* $p < .05$
** $p < .01$

Correlations ootained between scores from the observable performance factors from derived from the COKER and scores from the Core Battery of the NTE appeared to have little meaning. A correlation of .35 ($p<.05$) was noted between Professional Knowledge and classroom participation. Also the correlation of 0.35 ($p<.05$) between Communication Skills and presenting materials at various cognitive levels may have some meaning. Assuming that the COKER is a reliable and valid instrument for measuring classroom performance the use of the NTE in predicting success in the classroom was questioned.

Several significant correlations ($p<.05$) were noted between principals' overall ratings of performance and the variables of focus of the lesson, organization of presentations, class participation, evaluation, organization of the classroom, classroom management, and presentation of materials. However, Coker, Medley and Soar (1980), have questioned the validity and reliability of principals ratings of teachers performance.

KOI Results

Data were available from a sample of 38 individuals who had completed the bachelor's degree and were in their first year of teaching. Mean scores achieved on the PPST and the Core Battery of the NTE were as follows: PPST Reading (READ) 174, SD=4; PPST Writing (WRITE) 175, SD=6; PPST Mathematics (MATH) 175, SD=3; NTE Communication Skills (CS) 660, SD=8; General Knowledge (GK) 657, SD=10; and NTE Professional Knowledge (PK) 659, SD=11. The overall ratings of performance of the subjects by the building principals was also used in this study. The mean score on a five point scale (1=low to 5=high) was 4.1 (SD=0.7). Table 4 shows a summary of the correlations between the 14 factors from the KOI and the PPST sub-test scores, NTE Core Battery scores and the overall principals rating.

Several variables from the KOI were correlated significantly with scores achieved by students on the PPST. However, there is no logical explanation for these results. For example why should should there be a high positive relationship between classroom management and achievement on a reading test and a mathematics test.

An examination of the correlation pattern between the NTE scores and the KOI indicated only two significant correlations that might have an explanation. It was noted that there was a correlation of 0.39 ($p<.05$) between the General Knowledge test of the NTE and the classroom factor of questioning. Also significant correlations were noted between classroom involvement and the General Knowledge ($r=0.37$, $p<.05$) and Professional Knowledge ($r=0.46$, $p<.01$) test from the NTE. These seemed probable knowing the nature of the classrooms and the type information that was covered on the NTE.

Significant correlations were noted between principals' overall ratings and the KOI factors of questioning, presentation of new content, classroom involvement, and enthusiasm. Again, based on the literature these relationships could be questioned.

Table 4

Product Moment Correlations of Scores from the KOI, the NTE, PPST, and Overall Principals' Ratings (N=29)[1]

KOI Variables	READ	WRITE	MATH	CS	GK	PK	Prin.
Questions	-19	-07	10	23	39*	28	50**
Responses	13	16	51**	-00	-01	22	23
Praise	05	-24	-10	-05	-05	-40	06
Communications	00	00	00	19	-31	-33	-02
Management	62**	-06	57**	01	12	00	-05
Time Manage.	34*	14	21	09	18	05	-27
Seatwork	-07	-06	05	-13	-55**	-27	02
Begins Prom.	-22	15	16	-06	08	-13	-22
Academic Focus	-11	-04	26	-04	-24	-17	13
New Content	27	-01	16	11	08	26	46**
Questioning	-50**	-08	-21	-31	-27	-04	23
Directions	16	28	30*	18	-05	05	30
Involved	37*	22	17	27	37*	46**	47**
Enthusiastic	08	-14	-18	03	15	-10	51**

[1]Decimal points omitted.
* $p < .05$
** $p < .01$

Implications of the Study

The results of the study have implications for the improvement of teacher education programs. Generally the results of administering the PPST or the NTE provided little information that will be useful in predicting in a quantitative manner the success of an individual after they graduate and enter the classroom. Although there was little relationship between the NTE or the PPST and the observation systems utilized, these instruments still provided information useful in determining student deficiencies in potential basic skills areas. The PPST and NTE should not become, respectively, as they have in several states the sole criteria for admission to study in professional education and licensure as a teacher in the classroom.

This study should be replicated with a larger sample of subjects. Other studies needed to be initiated to determine the best predictors of success in the classroom. Only in this way can students be guided into proper careers during college and in turn the teaching profession will hopefully be improved.

References

Amidon, E. J. & Flanders, N. A. (1971). The role of the teacher in the classroom. Minneapolis: Association for Productive Teaching.

Ayers, J. B. (1988). Student assessment of teaching performance and classroom performance. Presented at the meeting of the Mid-South Educational Research Association, Louisville.

Ayers, J. B. (1987). Another look at the concurrent and predictive validity of the National Teacher Examinations. The Journal of Educational Research, 81(3), 133-137.

Ayers, J. B., & Qualls, G. S. (1979). Concurrent and predictive validity of the National Teacher Examinations. The Journal of Educational Research, 73(2), 86-92.

Coker, H., Medley, D. M., & Soar, R. S. (1980). How valid are expert opinions about effective teaching? Phi Delta Kappan, 62(2), 131-134.

Coker, H., & Coker J. G. (1982). Classroom Observations Keyed for Effectiveness Research-User's Manual. Atlanta: Georgia State University/Carroll County Teacher Corps Project.

Rudner, L. M. (1988). Teacher testing-an update. Educational Measurement: Issues and Practices, 7(1), 16-19.

Ryans, D. G. (1960). Characteristics of teachers. Washington: American Council on Education.

Siebert, P. Christine (1989). Relationships between the ACT, PPST, NTE, ACTCOMP and Achievement as Measured by College Grade Point Averages. Unpublished manuscript, Tennessee Technological University, Center for Teacher Education Evaluation, Cookeville, TN.

Tuckman, B. W. (1976). The Tuckman Teacher Feedback Form (TTFF). Journal of Educational Measurement, 13, 233-237.

Western Kentucky University. (1987). Manual for observing teacher behavior with the Kentucky Career Ladder Classroom Observation Instrument. Bowling Green, KY: Western Kentucky University, College of Education.

Chapter V

Relationships Between Scores from the ACT, PPST, NTE, ACTCOMP, and GPA

Paula C. Sibert

Students enrolled in teacher education programs are being required to complete more and more tests. For example in Tennessee students must complete the American College Test (ACT) prior to admission to a public institution. Prior to admission to professional education programs they must complete the Pre-Professional Skills Tests (PPST). At the time of completion of their programs of study they must complete both the College Outcomes Measures Project test of the American College Testing program (ACTCOMP) and the National Teacher Examinations (NTE) and have accumulated (in many cases) an overall grade point average of at least 2.50. Is all of this testing really necessary?

Previous studies have shown that scores on the PPST can be predicted from scores achieved on the ACT at the time of entrance to college (Stoker & Tarrab, 1985; Aksamit, Mitchell, Pozehl, 1987; Heard & Ayers, 1988). Jones (1982) concluded that scores on the NTE could be predicted from ACT scores. The purposes of the present study to examine the relationships between scores from the ACT, PPST, NTE, ACTCOMP and GPA and to explore the use of several variables in predicting scores on the outcomes measures.

Method

The subjects for this study consisted of 615 students who were graduates of two institutions of higher education in Tennessee. There were no significant differences in scores between the subjects from the two institutions on the four tests and overall GPA. All subjects had completed the three tests of the NTE Core Battery (Communications, General Knowledge and Professional Knowledge) and the ACTCOMP between the Fall of 1985 and the Summer of 1988. The ACT was completed prior to admission to the institution and the three tests of the PPST (Reading, Mathematics, and Writing) were completed, usually at the end of the sophomore year in college. Overall GPAS were taken from transcripts. Descriptive statistics were computed for all variables and regression analysis was used to predict the outcomes measures of the NTE, ACTCOMP, and GPA.

Results

The means, standard deviations, and correlations of all variables were presented in Table 1. Mean scores for the subjects indicated they had achieved between the 50th and 55th percentile (national morms) on the four tests. Correlations significant beyond the .01 level were noted between all variables in the study.

An examination of the relationships between the ACT and the PPST indicated, that in general, the ACT Composite score was the best predictor of

scores on the three tests of the PPST. These findings are in agreement with those previously reported (Stoker & Tarrab, 1985; Aksamit, Mitchell, Pozehl, 1987; Heard & Ayers, 1988).

Table 1

Means, Standard Deviations and Correlations of All Variables[a]

			Correlations											
	$\bar{X}$	SD	2	3	4	5	6	7	8	9	10	11	12	13
1. ACT English	20	5	55	58	59	80	66	39	64	61	71	63	61	54
2. ACT Mathematics	17	7		48	58	81	46	70	41	51	52	67	48	42
3. ACT Social Sciences	18	7			68	83	66	42	47	72	65	72	63	32
4. ACT Natural Science	21	6				86	61	59	44	61	60	70	55	40
5. ACT Composite Score	19	5					74	66	59	74	73	80	67	54
6. PPST Reading	177	5						57	57	70	76	76	74	47
7. PPST Mathematics	178	6							40	57	55	75	54	38
8. PPST Writing	176	3								49	62	55	51	43
9. ACTCOMP Total Score	186	16									56	62	58	37
10. NTE Communications	661	10										76	76	54
11. NTE General Knowledge	659	11											71	47
12. NTE Professional Knowledge	660	10												57
13. Overall GPA	3.09	0.44												

[a]Decimal points were omitted. All correlations were significant beyond the .01 level.

The best predictor of performance on the ACTCOMP was the ACT Composite score (r=.74). The best overall predictors of success on the three tests from the NTE was the Reading test score from the PPST and/or the ACT composite score. It should be kept in mind that the PPST and the NTE are closely related tests produced by the same organization. The best predictor of GPA was the ACT Composite score or the ACT English score (r=.54).

Stepwise multiple regression equations to predict success on the ACTCOMP, each of the three tests of the NTE, and GPA were developed using scores from

the ACT and the PPST. Variables were entered into the equation if they increased R^2 beyond the .05 level of significance. The ACT composite score entered at a significance level of .05 or less in the prediction of all of the scores except GPA.

Almost 55% of the variance was accounted for in the prediction of scores on the ACTCOMP by the ACT composite score. The addition of the ACT science and PPST reading scores accounted for an additional 13% of the variance. PPST reading scores accounted for 55% of the variance in the prediction of scores on the communications test of the NTE. Addition of the ACT English test and the ACT composite score accounted for an additional 15% of the variance. The ACT composite score accounted for 64% of the variance in the prediction of scores on the general knowledge test of the NTE. Addition of the PPST reading and mathematics tests scores increased the total variance accounted for to 77%. The best predictor of scores on the professional knowledge test of the NTE was the PPST reading score which accounted for 55% of the variance. Additions of the ACT science and Act composite scores increased the variance an additional 5%. The total variance accounted for in the regression equation to predict GPA was 27%. A combination of the PPST mathematics and ACT English scores were the best predictors. Following are the regression equations for the prediction of the outcomes measures.

ACTCOMP Total Score	= - 7.12 + 0.90(PPST Read) - 0.68(ACT Sci) + 2.50(ACT Comp)
NTE Communications	= 525 + 0.65(PPST Read) + 0.61(ACT Eng) + 0.46(ACT Comp)
NTE General Knowledge	= 451 + 0.45(PPST Read) + 0.61(PPST Math) + 1.11(ACT Comp)
NTE Professional Knowledge	= 494 + 0.84(PPST Read) - 0.39(ACT Sci) + 1.31(ACT Comp)
Overall GPA	= -0.38 + .0155(PPST Math) + 0.371(ACT Eng)

The R^2 for each of the equations was, respectively, ACTCOMP = 0.68, NTE Communications = 0.70, NTE General Knowledge = 0.77, NTE Professional Knowledge = 0.60, and Overall GPA = 0.27.

Conclusions

The empirical evidence from this study indicated the ACT composite score was the overall best predictor of success on the three tests of the PPST. Further the ACT composite score or the PPST reading score were the best predictors of success on the ACTCOMP and the three tests of the NTE Core Battery. Combinations of scores derived from both the ACT and the PPST improved the prediction of achievement on the standardized outcomes measures. Prediction of achievement as measured by the GPA, utilizing the ACT and PPST as predictors was questionable. The results neither supported or refuted the use of any of the tests. The evidence from the study indicated, however, that

scores from the ACT were a reasonable predictor of success on the outcomes measures. It appeared questionable whether there is a need for all of the testing that has become a part of teacher education programs.

References

Aksamit, D., Mitchell, J. V., & Pozehl, B. J. (1987). Relationships between PPST and ACT scores and their implications for the basic skills testing of prospective teachers. Journal of Teacher Education. 38(6), 48-52.

Heard, S. A., & Ayers, J. B. (1988). Validity of the American College Test in predicting success on the Pre-Professional Skills Test. Educational and Psychological Measurement. 48, 197-200.

Jones, J. W. (1982). The relationship between prospective teachers'scores on the American College Test and their National Teacher Examinations scores in selected groups of participants (Doctoral dissertation, East Texas State University, 1982). Dissertation Abstracts International, 43, 1369A-1370A.

Stoker, W. M., & Tarrab, M. (1985). The relationship between Pre-Professional Skills Tests and the American College Test. Teacher Education and Practice. 2(1), 43-45.

Chapter VI

Authoritarianism in Beginning Teachers

Lori A. Birdwell

Authoritarianism in teachers has been the subject of a variety of studies. For example McGee (1955) and Getzels and Jackson (1963) have reported studies utilizing the California F-scale (Adorno, et al., 1950) in examining the relationship between levels of exhibited authoritarianism and the effectiveness of teachers in the classroom. They concluded that teachers who tended to be non-authoritarian in their beliefs were better teachers. Authoritarianism has been linked to negative characteristics in teachers. Surlin (1976) found that high authoritarian college students indicated less concern for ethics and social responsibility than low authoritarian students. Harris, et al. (1982) found that authoritarian educators were more shy reserved, tense, frustrated, apprehensive, and unlikely to compromise, than less authoritarian educators. Ayers (1983) concluded that elementary teachers who tended to be less authoritarian in their believes were more effective in the classroom.

In related studies Tom (1984) found that teachers who tended to be authoritarian in their beliefs had higher grade expectations for middle class students than for lower class students, and higher expectations for occupational attainment for white females than for Asian females. Ayers, Rohr and Rohr (1977) found that subjects with more authoritarian beliefs perceived women to belong to a more subordinate role. All of these studies have shown that authoritarianism is related to negative characteristics and behaviors, many of which can be linked to whether an individual will be successful in the classroom.

Purpose

The purpose of this study was to examine the level of anti-democratic tendencies (authoritarianism) exhibited by groups of beginning teachers, who were graduates of one institution, as measured by the California F-scale Forms 45 and 40. The study, in part, replicated the work of earlier individuals and brings a new perspective on teacher preparation programs in the 1980s. More specifically the study was designed to test the following null hypotheses of no significant difference at the .05 level between F-scale scores for various groups.

1. There was no significant difference in F-scale scores between males and females.

2. There was no relationship between F-scale scores and age.

3. There was no significant difference in F-scale scores between subjects who had completed the bachelor's or master's degree.

4. There was no significant relationship in F-scale scores and year of graduation.

Method

Each year from 1973 through 1988, a sample of graduates (N=602) of the bachelors' and masters' degree programs for the preparation of teachers at Tennessee Technological University, were asked to participate in a longitudinal study of their performance after entering teaching. During the Spring of their first full year of teaching each graduate in the study completed the California F-scale, Forms 45 and 40. The 28-item instrument is related to opinions regarding a number of social groups and issues about which some people agree and others disagree. The subjects were asked to respond to each statement on a scale ranging from 1 to 7 (midpoint = 4) representing opposition (disagreement) to strong support (agreement). Total scores on the instrument can range from 28 to 192, with a mid-point of 112. Higher scores indicated subjects who were more authoritarian in their beliefs, and vice versa. Age, sex, and year and degree completed were taken from University records. Descriptive statistics, including Pearson Product Moment correlations were computed for the variables and the t-test and the analysis of variance technique (ANOVA) were used to test the hypotheses.

Results

The mean F-scale score for the total group was 102.6 (SD=21.0). The mean score for the 168 males was 106.9 (SD=21.0), while the mean score for the 434 females was 100.9 (SD=20.8). Application of the t-test indicated there was a significant difference between the two groups ($t=3.15$, $df=600$, $p<.05$). The males tended to be more authoritarian in their believes. The null hypothesis of no significant difference between males and females was rejected. This was in agreement with the findings of Ayers and Qualls (1982).

A Pearson Product Moment correlation was computed between F-scale scores and chronological age of each subject. There was no relationship between the two variables ($r=.04$, n.s.). The subjects were divided into three age groups. Group 1 included 241 subjects aged 21 through 25 which represented traditional college-age students at the time of the completion of their degrees. Group 2 included 237 subjects aged 26-36 and represented graduates who had already been in the world of work and had developed an independent life style. Group 3 included 124 subjects aged 36-59 who were already established in life and work. The mean scores for the three groups were, respectively, 101.6 (SD=21.4), 103.6 (SD=20.8), and 102.5 (SD=20.5). Application of the ANOVA indicated that there were no significant difference between the three groups ($F=0.55$, $df=2/599$). The hypothesis of no relationship between age and F-scale score was retained.

An examination was made of the difference in F-scale scores between those who held only the bachelor's degree (N=333) and those who held the master's degree (N=269). The means for the two groups were, respectively, 102.4 (SD=21.3) and 102.8 (SD=20.6). Application of the t-test indicated there was no significant difference between the two groups ($t=0.20$, $df=600$). The null hypothesis of no significant difference between bachelor's and master's level graduates was retained.

The relationship between year of graduation and F-scale score was examined by computing a Pearson Product Moment correlation between the two variables. There was a small but significant relationship between the two variables ($r=.14$, $p<.01$). In order to further examine the relationship and determine differences in the subjects, it was reasoned changes had occurred in society the 15 year period that could have affected the subjects. An examination of the data indicated that differences in scores appeared to be related to each other in three distinct groups. The groups included subjects who completed their degrees between 1973 and 1975 (Vietnam era), 1976 and 1980 (post Vietnam era), and after 1980 (era of conservatism). Table 1 shows a summary of the descriptive statistics for the three groups. F-scale scores increased over the three periods, while the ranges and standard errors of measurement of the means decreased. The groups became more authoritarian in their beliefs. Application of the ANOVA indicated there were significant differences between the groups ($F=3.64$, $df=2/599$, $p<.05$). Application of the t-test indicated there was a significant difference between the 1972-75 and 1981-88 groups. The null hypothesis of no significant relationship between F-scale scores and year of graduation was rejected.

Table 1

Descriptive Statistics for F-scale Scores by Year Group

Group	N	$\bar{X}$	SD	Range	Std. Error
1973-1975	82	97.3	23.5	44-177	2.6
1976-1980	202	101.7	21.2	45-165	1.5
1981-1988	318	104.4	20.0	46-160	1.1

Conclusions

Previous research has shown that teachers who are more open and less authoritarian in their beliefs may be more effective in the classroom. The present study indicated that a trend may be developing. Either individuals who are more authoritarian in their beliefs are entering the teaching profession or college students as a whole are becoming more authoritarian due to the changes in society.

In either case, action should be taken to overcome these changes. It has been shown that teachers who are less authoritarian tend to be better teachers. Therefore, teacher education programs need to take this into account through appropriate counseling and testing programs for perspective teachers. In turn, it was felt that teacher education programs can be improved and the end result will be teachers that are more attuned to society and working with children.

References

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., Sanford, R. N., Aron, B., Levinson, M. H., & Morrow, W. (1950). The authoritarian personality. New York: Harper.

Ayers, J. B. (1983, November) Authoritarianism and teacher effectiveness. Paper presented at the annual meeting of the Mid-South Educational Research Association, Nashville, TN.

Ayers, J. B., & Qualls, G. S. (1982, Fall). Authoritarianism and teachers' classroom behavior. Reading Improvement, 19(3), 219-222.

Ayers, J. B., Rohr, M. E., & Rohr, E. M. (1977, November). Attitudes toward women, college major, and authoritarianism. Paper presented at the annual meeting of the Mid-South Educational Research Association, Birmingham, AL.

Getzels, J. S., & Jackson, P. W. (1963). The teachers personality and characteristics. In N. L. Gage (Ed.), Handbook of research on teaching (pp. 522-23). Chicago: Rand McNally.

Harris, K. R., Halpin, Glennelle, Halpin, Gerald, & Roach, D. (1982, May). The relationships between pupil control ideology, self-concept, and teacher personality: Dimensions of teacher effectiveness. Paper presented at the annual meeting of the International Communication Association, Boston, MA.

McGee, H. M. (1955). Measurement of authoritarianism and its relation to teachers' classroom behavior. Genetic Psychology Monographs, 52, 89-146.

Surlin, S. H. (1976). Fatalism and authoritarianism: Predictors of professional attitudes in journalism. Journalism Quarterly, 53, 68-73.

Tom, D. Y. (1984). Influences of student background and teachers authoritarianism on teacher expectations. Journal of Educational Psychology, 76(2), 259-265.